

PLAN-CADRE DE COURS		
Titre du cours : Biostatistique		
Numéro du cours : 201-720-SL	Pondération : 2-2-2	Unités : 2,00
Place du cours dans le programme : 4 <sup>e</sup> session		
Cours prérequis (s'il y a lieu) : Il est préférable de faire <i>Bioinformatique</i> (145-215-SL) avant.		
Objectifs du cours : Comprendre le vocabulaire statistique et appliquer certaines méthodes de travail et d'analyse statistique afin d'étudier les variations biologiques, les phénomènes qui s'y rattachent et les problèmes qui en découlent.		
Compétences : 043W et 044A		
Contexte de réalisation : - À partir de situations appliquées au domaine de la biologie. - À l'aide : <ul style="list-style-type: none"> <li>• de tables, de graphiques et de manuels de référence;</li> <li>• de logiciels appropriés, tel un chiffrier électronique.</li> </ul>		

### Compétence 043W

Établir le portrait d'une situation à l'aide de données statistiques.

Élément de compétence 1 : Effectuer la saisie informatisée des données.	
Critères de performance :	Objets d'apprentissage :
1. Formulation de requêtes adéquates en fonction des bases de données.	1.1 Utiliser des fonctions de base d'un chiffrier (révision de notions vues dans le cours <i>Bioinformatique</i> ): - addition, soustraction, division, multiplication, logarithme, puissance; - moyenne, écart-type, rang, racine carrée; - calcul du nombre de cellules contenant des valeurs numériques.  1.2 Utiliser des fonctions de base d'un chiffrier (nouveau contenu): - générer un nombre aléatoire; - rechercher selon un critère; - utiliser un filtre; - construire des tableaux de distribution; - construire des tableaux croisés dynamiques; - construire des graphiques (par exemple, histogramme, diagramme à bâtons, courbe de fréquences cumulées).
2. Utilisation adéquate des applications appropriées, incluant le formatage du chiffrier.	(voir ci-haut, critère de performance 1)
3. Saisie précise et complète des données.	3.1 Saisir et présenter les données adéquatement.

Élément de compétence 2 : Vérifier la validité des données fournies.	
Critères de performance :	Objets d'apprentissage :
1. Repérage des biais liés à la collecte des données.	1.1 Identifier les différentes sources de biais : - méthodes d'échantillonnage inadéquates; - erreurs de mesure; - instruments de mesure imprécis.

2. Évaluation juste du degré de précision et de pertinence des mesures fournies.	2.1 Distinguer une donnée exacte d'une donnée précise. 2.2 Prendre en considération l'approximation des données.
3. Rejet des données non valides et non pertinentes pour la poursuite des travaux.	3.1 Repérer les données expérimentales extrêmes parmi l'ensemble des données d'une expérience. 3.2 Déterminer et décrire le test à appliquer pour évaluer si une donnée (réplicat) est extrême : le Q-test (entre 3 et 30 données) ou méthode de l'étendue interquartile (si $n > 30$ ). 3.3 Appliquer le Q-test et la méthode de l'étendue interquartile pour évaluer si une donnée est extrême. 3.4 Discuter, en tenant compte du contexte de l'expérience, de si une donnée extrême peut être considérée comme étant aberrante et peut alors être rejetée d'une série de données (réplicats).

### Élément de compétence 3 : Déterminer la nature du traitement statistique nécessaire.

Critères de performance :	Objets d'apprentissage :
1. Caractérisation correcte des échantillons : types d'échantillons, types de paramètres et types de distributions.	1.1 Identifier unité statistique, population, échantillon et variable (dépendante et indépendante). 1.2 Identifier le type de données : qualitative (nominale ou ordinale), quantitative (discrète ou continue). 1.3 Identifier et appliquer les différentes méthodes d'échantillonnage : aléatoire simple, systématique, stratifié et par grappes.  <b>Mesures de tendance centrale et de dispersion</b> 1.4 Identifier et différencier les mesures de tendance centrale : moyenne arithmétique, médiane et mode. 1.5 Calculer et interpréter les mesures de tendance centrale. 1.6 Identifier et différencier les mesures de dispersion : étendue, étendue interquartile, variance (d'une population = $\sigma^2$ et d'un échantillon = $S^2$ ), coefficient de variation et écart-type. 1.7 Déterminer les quartiles $Q_1$ et $Q_3$ dans une série de données. 1.8 Calculer et interpréter les mesures de dispersion. 1.9 Nommer les avantages et les inconvénients des mesures de tendance centrale et des mesures de dispersion.  1.10 Définir et appliquer les notions de base des probabilités : expérience, espace échantillonnal, événement, définition classique d'une probabilité. 1.11 Calculer le degré de liberté (dl). 1.12 Décrire les distributions d'échantillonnage : - loi de Student (pour un échantillon plus petit que 30 ou $\sigma$ est inconnu); - loi normale (pour un échantillon plus grand que 30 ou $\sigma$ est connu); - loi binomiale (brièvement); - loi du Khi-carré. 1.13 Calculer et interpréter les intervalles de confiance : estimation de la moyenne et d'une proportion (ou fréquence relative).
2. Sélection des traitements et des tests statistiques en fonction de la nature des données et de l'hypothèse de départ.	2.1 Décrire l'utilité et les conditions d'applications de : - la comparaison de moyennes (test t, analyse de variance ou Anova); - la comparaison de proportions (Khi-carré – $\chi^2$ ); - la corrélation et la régression linéaire simple. 2.2 Décrire le coefficient de corrélation ( $r$ ) et de détermination ( $r^2$ ). 2.3 Appliquer un test d'hypothèse : définir hypothèse nulle ( $H_0$ ), hypothèse(s) alternative(s) ( $H_1$ ), type d'erreur ( $\alpha$ et $\beta$ ), règle de décision. 2.4 À partir d'une situation concrète liée à la biologie, déterminer les hypothèses, les erreurs $\alpha$ et $\beta$ ainsi que le test à effectuer.

<b>Élément de compétence 4 : Effectuer les traitements statistiques prévus.</b>	
Critères de performance :	Objets d'apprentissage :
1. Sélection des fonctions statistiques appropriées.	(voir élément de compétence 3, critère de performance 2)
2. Sélection des variables pertinentes.	2.1 Sélectionner les variables pertinentes dans un tableau de données selon le caractère étudié d'une population.
3. Exécution adéquate des fonctions sélectionnées et exactitude des calculs.	3.1 Appliquer les tests d'hypothèse suivants : <ul style="list-style-type: none"> <li>- test sur une seule moyenne, test sur deux moyennes, test sur <math>x</math> moyennes (comparaison de moyennes : test <math>t</math> et Anova);</li> <li>- test sur une seule proportion, test sur deux proportions, test sur <math>x</math> proportions (comparaison de fréquences : Khi-carré – <math>\chi^2</math>);</li> <li>- corrélation et régression linéaire simple.</li> </ul> 3.2 Déterminer le coefficient de corrélation ( $r$ ) et de détermination ( $r^2$ ).

<b>Élément de compétence 5 : Assurer la mise en forme et la présentation graphique des données.</b>	
Critères de performance :	Objets d'apprentissage :
1. Sélection du mode de présentation graphique le plus approprié.	1.1 Différencier les types de graphiques : <ul style="list-style-type: none"> <li>- diagramme en bâtons;</li> <li>- diagramme à bandes;</li> <li>- histogramme;</li> <li>- polygone de fréquences;</li> <li>- nuages de points.</li> </ul> 1.2 Choisir le type de graphique le plus approprié à une situation donnée.
2. Qualité des tableaux : <ul style="list-style-type: none"> <li>- pertinence des données présentes;</li> <li>- prise en compte des paramètres appropriés (structure, organisation et lisibilité facilitant l'interprétation).</li> </ul>	2.1 Respecter les normes du Guide méthodologique de rédaction d'un texte scientifique établi par le programme de bioécologie (vu en <i>Méthode scientifique</i> ) pour la présentation des tableaux.
3. Qualité des figures : <ul style="list-style-type: none"> <li>- présentation des points saillants;</li> <li>- conformité aux normes de présentation.</li> </ul>	3.1 Respecter les normes du Guide méthodologique de rédaction d'un texte scientifique établi par la le programme de bioécologie (vu en <i>Méthode scientifique</i> ) pour la présentation des figures.

<b>Élément de compétence 6 : Analyser et interpréter les résultats obtenus.</b>	
Critères de performance :	Objets d'apprentissage :
1. Définition correcte des différentes notions statistiques et de la terminologie utilisée.	(voir éléments de compétence 3 et 4)
2. Justesse des explications relatives aux traitements statistiques effectués.	(voir éléments de compétence 3 et 4)
3. Degré adéquat de précision et de signification des résultats présentés.	(voir éléments de compétence 3 et 4)
4. Justesse de l'interprétation compte tenu de la problématique de départ.	(voir éléments de compétence 3 et 4)

**Autres cours liés à cette compétence :** Aucun

## Compétence 044A

Utiliser les technologies numériques et informatiques dans le cadre de ses activités professionnelles.

### Élément de compétence 3 : Utiliser des outils informatisés de traitement statistique et de représentation de données.

Critères de performance :	Objets d'apprentissage :
4. Saisie adéquate des données fournies.	4.1 Compiler des données à l'aide d'un chiffrier.
5. Utilisation des fonctions de base relatives au traitement statistique de l'application.	5.1 Utiliser un chiffrier et un logiciel d'analyse de données pour toutes les fonctions statistiques vues dans ce cours.

**Autres cours liés à cette compétence :** *Bioinformatique* (145-215-SL)

#### Exigences particulières du département ou du programme (s'il y a lieu) :

Les exemples et les exercices doivent être appliqués à la biologie.

Respecter le Guide méthodologique de rédaction d'un texte scientifique (vu en *Méthode scientifique* (145-105-SL)).

Respecter la Politique départementale d'évaluation des apprentissages (PDÉA) du département de mathématiques.

#### Médiagraphie indicative, logiciels et/ou équipements suggérés :

##### Livres de référence

**Colin, Michèle et Gisèle Payette.** 2004. *Biostatistiques pour les techniques biologiques*. Publié par les auteurs. Sherbrooke.

**Grenon, Gilles et Suzanne Viau.** 1997. *Statistiques appliquées : initiation à l'analyse des données statistiques*. Gaëtan Morin éditeur.

**Howard, Christensen.** 1986. *La statistique : démarche pédagogique programmée*. Gaëtan Morin éditeur.

**Laliberté, Célyne.** 2005. *Probabilité et statistiques : de la conception à la compréhension*. ERPI.

**Mercier, André.** 1988. *Éléments de biométrie*. Les éditions Le Griffon d'argile.

**Ouellet, Gilles.** 1998. *Statistique et probabilité*. Les éditions Le Griffon d'argile.

**Scherrer, Bruno.** 1984. *Biostatistique*. Gaëtan Morin éditeur.

**Simard, Christiane.** 2002. *Notions de statistiques*. Les éditions Le Griffon d'argile.

Guide méthodologique: rédaction d'un texte scientifique.

##### Logiciels

Excel, SigmaPlot, Logiciel R, Statview, SAS, Systat, JMP, SPSS

#### Stratégies d'évaluation suggérées :

Au moins 3 examens théoriques (environ 75%)

Au moins 2 laboratoires (environ 25%)